

# The Mediaeval Blog. Towards a Database of 15th and 16th Century Dutch Dialect and Idiolect Data.<sup>1</sup>

Chris De Wulf, researcher at the Academy

## 1. A WELL-DOCUMENTED PERIOD IN THE HISTORY OF DUTCH: THE 13<sup>TH</sup> AND 14<sup>TH</sup> CENTURIES

The Gysseling Corpus of Middle Dutch texts up until the year 1300 in its digital format has given a real boost to methodical research of the history of Dutch. More specifically the volumes consisting of non-literary texts (predominantly charters) broke the mould of historical linguistics, and instigated a kind of diachronic research modelled on modern (data-driven) dialectology. For the first time, researchers would have access to a geographical diverse and stylistically fairly uniform (digital) database of Dutch varieties. This dialectological approach resulted in two dialect atlases of the 13<sup>th</sup> century:<sup>2</sup> Berteloot's *Bijdrage tot een klankatlas van het dertiende-eeuwse Middelnederlands*<sup>3</sup> and Mooijaart's *Atlas van Vroegmiddelnederlandse taalvarianten*.<sup>4</sup> The success of the so-called *Corpus Gysseling* (CG) has also led to a new, similar corpus of charters from the 14<sup>th</sup> century. Named after its initiators, the enriched *Corpus Van Reenen - Mulder* (CRM) contains close to 4000 charters, almost all of which have been lemmatised and POS-tagged.<sup>5</sup> Apart from several publications by the main participants in the corpus building process, the CRM resulted in a PhD on a corpus-fed localisation method,<sup>6</sup> and is currently providing data for a phonological

---

<sup>1</sup> I would like to thank Jozef Van Loon and Georges De Schutter for their useful remarks on a first draft of this paper.

<sup>2</sup> Although technically they predate the digitally available version of the corpus.

<sup>3</sup> Gent, 1984.

<sup>4</sup> Utrecht, 1993.

<sup>5</sup> The part within the CRM that has been contributed by the Royal Academy, has recently been published online. It consists of the C14NL-NeG Corpus, 441 charters from southern Flanders and western Brabant, published as XML transcripts, and the C14NL-PoS Corpus, 157 southern Flemish text (areas including Ieper, Oudenaarde, Zottegem, Velzeke) in digital (XML) transcripts, and with lemmas and PoS-tags added. See at <http://www.kanlt.be/ctb/corpus>. The C14NL has been designed as a regionally complementary addition to the original CRM and should be seen as part of it in its contents and the way it is organised. We will further use the name CRM for the entire integrated corpus, i.e. all 14<sup>th</sup> Century charters.

<sup>6</sup> Rem 2003. The compounders of the data are currently considering hosting options. Potential users that would like to obtain the data should contact Prof. Dr. van Reenen (Meertens Instituut, Amsterdam) or Dr. Rem (Radboud Universiteit, Nijmegen) to discuss temporary solutions.

atlas of 14<sup>th</sup> century Dutch.<sup>7</sup> Such research is possible because of 3 key requirements which the charters in the mentioned corpora meet. We refer to them as (1) Temporal and Spatial Localisation, (2) Low Social Status and (3) Originality.<sup>8</sup>

(1) Charters, and more specifically, the local charters that are seen fit for geolinguistic purposes, bare proof of legal transactions, commitments or judicial decisions. Their legal status originates from the honourable status of the people present<sup>9</sup> when the legal act occurred. Legal status could be proven by the fact that, apart from the parties that are concerned, these witnesses (usually aldermen who often acted as the issuers of the charter), confirm with their personal seal or with a joint seal of the institution they represent. They are usually enumerated at the beginning or the end of the charter, stating what alderman office (i.e. which locality) they belong to. The charter further mentions its date of issue. This means that all writing in any specific charter represents a certain time and place. The place of issue, however, can only be derived from text if only one place is mentioned. In other words, the issuer and the receiver of the legal text must reside in the same locality. If that is not the case, such a charter should not be considered to be part of the otherwise well-localised and well-dated corpus. Exceptions can be made, however, when only neighbouring localities occur in the same text, hence providing a small region as the place of origin. In that case, though, the ‘regional’ character of the charter should be mentioned in the charter’s metadata.<sup>10</sup>

(2) The requirement of Low Social Status is the historical linguist’s equivalent for the contemporary dialectologist’s Non-Educated Older Rural Male. However, no charters written by NORMs (apart from, quite obviously, the Male feature) are to be found as being able to write in the Middle Ages already implied (somewhat rare) education privileges. The Older

---

<sup>7</sup> De Wulf, PhD at the University of Gent, to be expected in 2009.

<sup>8</sup> These requirements have been derived from the strategy taken to build the CRM-database, which have been explained extensively by P. van Reenen & M. Mulder, ‘Een gegevensbank van 14de-eeuwse Middelnederlandse dialecten op computer’ in *Lexikos 3* (1993), p. 259-279. They have been repeated as an example of ‘good practices’ in extending the CRM, in P. van Reenen, ‘Het 14de-eeuwse Middelnederlandse oorkondencorpus als dynamisch-systematisch referentiekader voor taalkundig onderzoek’ in: Vanhoutte, E. (red.), *Talig Erfgoed. De zuidelijke Nederlanden in de 14de eeuw*. Gent, p. 59-72. The requirements were not followed by Gysseling when building his corpus, as he could afford to render all available charters of the 13th century. Gysseling often used the linguistic method for establishing certain charters’ place of origin. Later researchers using the Corpus Gysseling, however, had to exclude some of Gysseling’s material for methodological reasons (avoiding Gysseling’s *petitio principii*), and often did so only maintaining the text that met requirements similar to the ones we mention.

<sup>9</sup> Gentry, guild members and other people who held a certain office.

<sup>10</sup> In the CRM, this problem was assessed by adding a different format of *Kloeke* codes. As every charter got an exclusive reference number containing the date of origin and the locality coded in the Kloeke system (for example: I241p37501), the regional charters use a variation on the Kloeke code to not only show its regional status but also to show similarity to the most likely place of origin within that region, adding 500 units and thus tapping into unused numbers within the Kloeke system (I741r37501). Contexts on the Kloeke system can be found at <http://www.meertens.knaw.nl/projecten/mand/ECARTkartografie.html> (sic).

and Rural can best be rephrased as ‘a lack of mobility’, although, in most cases scribes and (worldly) clerks could have spent at least part of their training either away from home or with a non-local master. This problem is most apparent in religious orders and the like, that would draw in scribes from far beyond the borders of the local community. This has urged the constituters of at least the CRM to only consider charters that are linked to a worldly office, in most cases the local (and laic) aldermen administration.<sup>11</sup> Low Social Status in these matters means that the charters have been issued by and to one or more of these local administrations, and not, e.g. by a lord or some higher office. People of high social ranking, such as knights and most definitely counts and dukes, would draw upon a much larger geographical area for their clerks, hence not providing a local dialect in the writing issued from their offices.

(3) If, according to the above, we consider the writing in any single charter to be well localised in time and space, we can only do so because we assume charters do not contain layer upon layer of writing of different authors, spread over a vast amount of time. That assumption is partly defected, as most charters that have been preserved are direct copies of minutes taken at the time of the legal act described in the text. This means that charters can be copied by different scribes than the ones that took the original notes. We accept this small deviation in the assumption that charters represent one time and one place, provided we only take into account *so-called original* charters of which can reasonably be assumed that they came into existence shortly (presumably the same day or week) after the events described took place. That means that we must exclude copies that have most certainly been made at different times, by people distant from the original scribe. We consider the difference between what we call *originals* (which might be, in fact, ‘immediate copies’) and what we brand *copies*, key to our understanding of a corpus of charters as a group of texts that are well localised in space and in time.

Because of the abovementioned criteria, 13<sup>th</sup> and 14<sup>th</sup> century charters have proved useful hunting grounds for geographical and temporal variation in the mentioned eras. This is a time in which local language seems to prevail, especially in the writing of (local) clerks, provided that no distant or socially superior participants are involved. In terms of output and importance to the knowledge of Dutch dialects and Dutch language history, the

---

<sup>11</sup> The mediaeval equivalent of the local council, with administrative and judicial powers, led and executed by aldermen (Dutch: *schepenen*).

abovementioned databases are only equalled or surpassed by ‘modern’ day dialect databases from the late 19<sup>th</sup> and (mainly) 20<sup>th</sup> centuries.<sup>12</sup>

The gap that is left between the 14<sup>th</sup> and the 20<sup>th</sup> centuries, where vast dialect databases covering the entire Dutch speaking area are concerned, now remains to be filled. The 16<sup>th</sup> and 17<sup>th</sup> centuries are generally considered a pivotal period in which Modern Dutch has emerged. It was a politically rumbling time for the Low Countries, resulting in Holland’s cities drawing in scores of people from the South and the East of the Low Countries, and beyond. The era prior to these great changes, starting with the 15<sup>th</sup> century, can be seen as a reasonably static period in human traffic, possibly providing us with key data with regard to the regional origins of the melting pot that was to be New Dutch.

## 2. HOW TO FILL THE GAP

Since the 14<sup>th</sup> century, public and interpersonal administration (like the transactions we encounter in local charters) has been professionalised and centralised, gradually and at times abruptly, leading to present times of highly formalised written language, produced amidst and by legal and administrative professionals such as law practitioners and civil servants. This means that gradually in time, (databases of) charters dated after the 14<sup>th</sup> century will show us a type of formal written language that becomes more and more alienated from every day speech. Therefore we propose that historical linguists should ideally get hold of a double themed corpus: a (local) charter-based one for the 15<sup>th</sup> century, which ties in with the older databases on the one hand, and a corpus containing different forms of autobiographical writing, that seem to emerge in this exact period. We will call those *ego documents*. The word ego document is a Dutch term referring to autobiographical writing that includes diaries, memoirs, personal letters, (travel) journals and notes, but not (primarily) those that have been

---

<sup>12</sup> These are, of course, databases built on structured dialect queries, submitted in writing or, more recently, recorded and transcribed. The most recent ones with a wide geographical scope include the GTRP-data (morpho-phonological, not exclusively but most conveniently accessible at the Meertens Instituut), the SAND-data (syntactical, ib.), the regional dialect dictionaries (data yet to be made public, contexts at <http://fuzzy.arts.kuleuven.be/rewo>), and the somewhat older counterpart of all of these, the rather inaccessible RND (data is not digital but can be consulted in participating institutes such as the University of Gent - Dutch Ling. Dept., samples can be found on <http://www.let.rug.nl/~heeringa/dialectology/atlas/rnd/>). The real pioneer of dense geographical dialect inquiries, however, was Pieter Willems, whose late-nineteenth century data is available on microfilm in a limited amount of places (e.g. UGent, Dutch Ling. Dept.). His original notebooks are preserved at the Royal Academy of Dutch Language and Literature (KANTL, Gent). The KANTL library also hosts microfilm versions of these data, which will be digitalised in the near future thanks to funding provided by FWO-Vlaanderen.

written as an artistic expression, which we would rather classify as literature. Ego documents that have been written, as we might assume, for personal or family use, stand the best chance of rendering their authors' dialect or idiolect. We will further explain how we intend to build a database of this kind of writing, but it is worthwhile to give an overview of initiatives that help to close the gap between the 14<sup>th</sup> century and now, where collections of Dutch writing are concerned. As part of a corpus for linguistic purposes, however, fiction/literature cannot guarantee precise localisation and dating of the tokens found in it, not in the least because it usually involves editing and intertextuality to a degree that its main dialect or idiolect layer is heavily corrupted. Therefore, in a (historical) geographical linguist's definition, literature doesn't provide us with palpable original artefacts. Such considerations have already been put forward extensively in Van Loon (2001).

The following overview hence focuses on non-literary texts from the 15<sup>th</sup> and subsequent centuries.

### 3. WORK (BEING) DONE BY OTHERS FOR THE 15<sup>TH</sup> CENTURY AND ONWARD<sup>13</sup>

As there's a painful lack of structural funding for corpus building in the humanities in the Low Countries (and probably more so in Flanders than in the Netherlands), the existing initiatives for acquiring corpora of older Dutch are quite fragmented. Most of them have been initiated by individual researchers or research groups, and cater for their individual needs, as opposed to long running projects of data acquisition. Funding for broader scope corpus building and maintenance has sometimes successfully been acquired, however, by institutions such as the INL.<sup>14</sup> This does not erase the fact that, unfortunately, most attempts at building widely usable corpora are short-lived projects. They usually compete for funding against individual research, which, ironically, often depends on the envisaged corpora for their study material. The ongoing struggle for infrastructural needs hence impedes progress in innovative linguistic research.

---

<sup>13</sup> Some of the information below was provided by the Netwerk Diachronie, for which I would like to express my gratefulness.

<sup>14</sup> The Leiden-based Instituut voor Nederlandse Lexicologie, perpetually funded by the Nederlandse Taalunie.

Corpora that include non-literary Dutch texts from the late Middle Ages or later (15<sup>th</sup> up to the 19<sup>th</sup> century), are:

- *Cartago*.<sup>15</sup> Some 35000 charters from Groningen and Drenthe predating 1600 are visualised in facsimile and in (diplomatic) transcript (offered in HTML).
- *Non-Frisian Testaments* (NFT) from Friesland.<sup>16</sup> This database consists of diplomatic transcripts in earlier (printed) editions of Frisian Testaments written in Dutch from the 15<sup>th</sup> and 16<sup>th</sup> centuries are published as PDFs.
- *Corpus Ruesink* (CR)<sup>17</sup> or ‘Hofboek en Voluntaire protocollen Bredevoort’ contains charters (i.e. legal deeds) and accountancy regarding the former seignery of Bredevoort from the 16<sup>th</sup> and 17<sup>th</sup> centuries. The CR website offers an extensive inventory and transcripts of the (original) texts (as PDFs).

Projects that provide us with an (internet-based) inventory of texts, rather than the rendition of those texts, some of which might be classified as ego documents, include:

- *Narrative Sources*.<sup>18</sup> They provide an inventory of semi-fictional and non-fictional narrative texts originating in the Low Countries, spanning 10 centuries (6<sup>th</sup> - 16<sup>th</sup>), including chronicles, diaries, travel diaries, hagiographies etc. in Dutch, French and Latin. The database (i.e. the website) does not provide transcripts as such.
- *Sailing Letters*.<sup>19</sup> This initiative by the Dutch Royal Library (Koninklijke Bibliotheek Den Haag) lists 17<sup>th</sup> and 18<sup>th</sup> Century Dutch letters and accountancy that have been found in the (UK) National Archives. A small section of these letters is accessible in transcript (philological). Further exploration of this material is the subject of an NWO-project ‘Brieven als buit’ (M. Van der Wal) at Leiden University.<sup>20</sup>
- *www.egodocument.net*, created by Rudolf Dekker, is an extensive inventory and bibliography of ego documents (diaries and travel writing) in the Netherlands from the 1500s up until 1914.

Tightly knit with the post-mediaeval history of Dutch, is the history of creole languages with a Dutch base or under Dutch influence. Apart from the continental Dutch databases, information can be found in specialised corpora for these creoles: *The Suriname*

---

<sup>15</sup> Available at <http://www.cartago.nl>.

<sup>16</sup> At <http://www.fa.knaw.nl/fa/downloads/Geschiedenis,%20Letterkunde%20en%20Naamkunde/friese-testamenten-tot-1550/friese-testamenten-tot-1550-in-pdf>

<sup>17</sup> Contexts, inventory and transcripts at <http://www.heerlijkheidbredevoort.nl>.

<sup>18</sup> Cf. [www.narrative-sources.be](http://www.narrative-sources.be) (in English and in Dutch).

<sup>19</sup> Contexts and the actual inventory can be found on the website: <http://www.kb.nl/sl/index.html>.

<sup>20</sup> Cf. <http://www.brievenalsbuit.nl>.

*Creole Archive* (SUCA),<sup>21</sup> *Negerhollands Creole Archive* (NECA)<sup>22</sup> and the *Cape Pidgins Archive* (CPA).<sup>23</sup>

Projects for other languages include the above-mentioned Narrative Sources website (Low Countries) and the France-based *Egodoc* project,<sup>24</sup> which aims at rendering editions of 18<sup>th</sup> and 19<sup>th</sup> century ego documents.

#### 4. EDITIONS OF AUTOGRAPHICAL MIDDLE DUTCH EGO DOCUMENTS AS SOURCES FOR HISTORICAL DIALECT AND IDIOLECT GRAMMARS.<sup>25</sup>

The CTB is currently working on a corpus containing 15<sup>th</sup> and 16<sup>th</sup> century ego documents, hence aiming to meet part of Dutch historical linguists' requirements that we have stated above. In the first stage of project, we intend to reflect on the selection and edition process, which should result in one finished digital edition, a set of 'good practices' for the whole process, and an (online) inventory of manuscripts eligible to be rendered in the future. We will discuss the selection criteria we use.

In Van Loon (2001), the case of two types of non-literary texts has been put forward. Apart from ego documents, he also considers judicial proceedings, and, to be more precise, *witness accounts*, of possible interest. Such written representation of direct speech (*oratio obliqua*) is as close as we can get to the idiolect and dialect of people from centuries past. Van Loon mentions Vanacker (1963), a study involving witness accounts from the 15<sup>th</sup> to 17<sup>th</sup> centuries, as an example of the potential of these judicial proceedings. However, Van Loon also unveils restrictions in the use of such documents, namely the lack of thematic diversity: "Vermeldingen van wapens, lichaamsdelen, eigendomskwesties e.d. voeren de boventoon" (Van Loon 2001: 421). This problem also occurs when working with charters. Ego documents and especially diaries, are more promising as the genre does not entail such thematical

---

<sup>21</sup> See [www.ru.nl/suca](http://www.ru.nl/suca).

<sup>22</sup> See Cefas van Rossem, Hein van der Voort, *Die Creol taal. 250 years of Negerhollands texts*. Amsterdam 1996.

<sup>23</sup> Not published.

<sup>24</sup> Contexts in: Pierre-Yves Beaurepaire, Dominique Taurisson (éd.), *Actes du colloque international de Montpellier 'Les ego-documents à l'heure de l'électronique. Nouvelles approches des espaces et des réseaux relationnelles*. 2003, Presses universitaires de Montpellier, Montpellier. The data, which for now exists of the fac simile and transcript of the journal of knight *Marie Daniel Bourrée de Corberon*, can be found at: <http://arcanews.univ-montp3.fr/egodoc>.

<sup>25</sup> The scope of this project has also been revealed (in Dutch) at <http://www.kantl.be/ctb/project/2008/egodocumenten.htm>. Supervisors are Prof. Dr. J. Van Loon and Prof. Dr. F. Willaert (University of Antwerp and members of the Royal Academy of Dutch Language and Literature). The project is being executed by the author of this paper.

restrictions. Moreover, the three 3 key requirements that we saw met in the mediaeval charter corpora, and that have proven useful (or even necessary) for historical (geo)linguistic research, can easily be applied to (certain types of) ego documents.

(1) Temporal and Spatial Localisation. Everyday autobiographical texts such as diaries, travel journals and family accounts typically provide entry dates. If they have been kept as logbooks (as opposed to memoirs), the entry date coincides with the time of writing. Many of the diary type texts we find, however, mix different styles of writing (memoirs, dead and live chronicle, log-style entries).<sup>26</sup>

As the narrative of memoirs seems to end at the time of writing, those dates provide temporal beacons for the language found, as well as (dated) diary-style entries.

The geographical localisation, or, in other words, pinpointing a dialect to an ego document, is a different matter. It depends on the history of the author whether they can be considered representative for a certain local or regional language. Therefore, it is paramount to have additional information on the author's life, their place of birth, places of residence, professional activities and the people they were surrounded with.<sup>27</sup> Most of the information we need should be provided for by the author himself, exactly because of the nature of their writing. This implies that non-autographical texts, or writings by anonymous authors remain beyond the scope of our well-dated and well-localised corpus.<sup>28</sup>

(2) Low Social Status. As stated before, literacy would have been out of most people's reach in the Middle Ages. The people that left us their writing would have had a certain social status that co-occurs with geographical and social mobility. Their social status might thus be a corrupting factor in their rendering of local language. Many of the ego documents that can be found in manuscript today, have been made by people who have a professional tendency for keeping track of accounts, such as parish priests, merchants and diplomats.<sup>29</sup> If they seem to have lacked further artistic ambition, this will raise the chances of encountering a very personal (idiolect or dialect) form of writing. This would have to be assessed on an individual

---

<sup>26</sup> *Jan De Pottre* (°1525-<sup>+</sup>1601) combines a dead chronicle with his memoirs, *Simon de Rikelike* (unknown-c. 1601) states his accounts (in Dutch) and adds a chronicle (in French), *Pieters Segers* (1551-1621) writes his memoirs in one go, and subsequently diary entries for successive years.

<sup>27</sup> For example, both of the author's parents should preferably be locals of the same area.

<sup>28</sup> Two well-known interlinked examples of these are the travel journals of brothers *Karel* (1533-1584) and *Lieven* (unknown-1610) *Rijm*. Their travel journals have (most probably) been written by people from their entourage. Good guesses are a certain *Johannes van den Brande* and *Pieter van den Doerpe* respectively, but other than the fact that they worked for (or at least travelled with) these brothers, we know nothing about them.

<sup>29</sup> Limburgian diarist *Christiaan Munters* was a priest, Brabantine *Jan De Pottre* was in commerce, as were *Simon de Rikelike* (Bruges), *Herman Pottey* (Brabant) and *Pieter Segers* (Gent). *Edmond van Dynter* (Brabant) was a professional writer, however, as annalist and secretary to the Burgundian duke.



basis, making an informed decision on whether a certain author intended his writing to be read by a wider audience than himself or his family, for instance.

(3) Originality. As being able to write also implies being able to read, whoever felt the urge of describing their lives and environments, could very easily be influenced by other writings of a similar nature. Diaries that describe events of a non-local nature might include rephrased or even just copied parts of other writing such as dead chronicles.<sup>30</sup> Within manuscripts that combine those, the constituting parts have different status in terms of originality, and must thus be treated differently.

The more obvious concept of originality defined by the *autographical status* of a text is, of course, equally paramount. As goes for charters, we should only involve the original manuscript in our corpus, and not, for example, a copy that has been made centuries later. As there has been a surge of interest in ego documents in the 19th century, many older manuscripts have been printed. These editions can add valuable information when contextualising the original,<sup>31</sup> but as editing principles were very poor compared to those of modern philologists', access to the original manuscript remains necessary.

## 5. CASE STUDY: THE VOWEL OF *BRACHT* IN THE BRUSSELS DIALECT OF JAN DE POTTRE

### 5.1. INTRODUCTION: TWO TYPES OF VELARISATION

One of the discriminating orthographical and phonological markers of the Dutch dialects is the velar vowel that may occur in Dutch *bracht*, *brachten* and *gebracht*<sup>32</sup> (Eng. 'brought', the singular and plural simple past, and the past participle of the verb 'to bring' respectively).

Germanic /ǣ/, followed by a cluster consisting of a nasal, a fricative and a stop, would be susceptible to lengthening and nasalisation (combined with regression of the nasal consonant).<sup>33</sup> In a secondary stage, this vowel could easily lose its nasal quality and shorten again. For example: brǣŋχt > brǣ̃χt > brāχt > brǣχt.<sup>34</sup> The same process led to Dutch *zacht* ('soft'). Van Loon (1986: 28-29) assumes that the originally medial a in this position, could

---

<sup>30</sup> *Jan De Pottre's* manuscript, for instance, consists of two distinct parts: the first being a chronicle covering the 14<sup>th</sup>, 15<sup>th</sup> and the first part of the 16<sup>th</sup> century, the second part is a mixture of his memoirs and local events during his lifetime, that is more likely to have been witnessed by him (although not written on a day by day basis).

<sup>31</sup> And to facilitate transcribing the original.

<sup>32</sup> We will refer to all three combined as 'bracht'.

<sup>33</sup> As pointed out by Schönfeld (1970: 29) and FAND (I: 16).

<sup>34</sup> Cf. FAND (ib.).

have had a fairly velar allophone, influenced by the following nasal consonant, which would have brought on a realisation such as  $br\bar{o}\chi t$  in the second stadium. The allophone lost its allophonous status after total nasal loss, and would have evolved differently in the east and west of the language continuum. It would have merged with its closest neighbour. In the east, that was (long)  $\bar{a}$ -type vowel (the representative of Germanic  $\hat{e}$ ), bringing on the evolution, as described above, resulting in  $br\check{a}\chi t$ . The coastal dialects lacked such a long  $\bar{a}$ , and the vowel in *bracht* could only merge with  $\bar{o}$ . Note that English (too) has a similar o-type vowel in *brought* and *soft*. The western Dutch equivalents are *brocht* and *zocht*. The trajectory proposed here, would be:

$br\check{a}\chi t$	>	$br\bar{o}\chi t$	>	$br\bar{a}\chi t$	>	$br\check{a}\chi t$ (south eastern dialects)
			>	$br\bar{o}\chi t$	>	$br\check{o}\chi t$ (western dialects)

While velar realisation in *bracht*, *zacht* etc. (~current short /a/ in modern Standard Dutch) is still a coastal dialect property to this day, the more continental dialects of Brabant generally have velar realisations of what is now long /a:/ in modern Standard Dutch: long West Germanic / $\hat{a}$ / and lengthened (originally short) W.G. / $\check{a}$ / (when in syllable-final position). Examples are Brabantine *st[ɔ:]n* (D. ‘staan’, ‘to stand’), *m[ɔ:]ke* (D. ‘maken’, ‘to make’). The impact of this expansive Brabantine phenomenon is shown in map 1, portraying velar vs. open/medial vowels in Dutch *maken*. It is based on modern dialect data.<sup>35</sup> The inlay map shows the density of the material: darker patches have more data; brighter patches are less reliable and can best be ignored. In the main map, dark colouring represents open/medial vowels (such as /a./, lighter colouring is caused by occurrences of velar vowels (such as /ɔ./). Note that the velar vowel quality in Brabant, Limburg and Groningen (North East) are probably three different phenomena.<sup>36</sup> The velar vowel in the east of Flanders, however, has most likely been caused by influence from adjacent Brabant.

<sup>35</sup> GTRP (Goeman-Taeldeman-Van Reenen Project): 20<sup>th</sup> Century phonological and morphological data. References at <http://www.meertens.knaw.nl/mand/database/> (in Dutch). In the data used for maps 1 and 3, [ɒ] to [u] span the range of velar vowel quality, [q] to [ɪ] are open/medial. The cut-off point, effectively, lies between [ɒ] and [q]. The mapping technology used for the three maps in this article has been provided by Evert Wattel. Contexts in Wattel & Van Reenen (1995).

<sup>36</sup> Limburg velarisation of / $\bar{a}$ / is definitely older than the Brabantine kind. We encountered the Limburg form as early as the 13<sup>th</sup> and 14<sup>th</sup> century, when there is still no sign of the Brabantine counterpart. It is very unlikely that the Limburg vowel would have been imported in Brabant, given the peripheral status of Limburg. Cf. De Wulf 2009 (to be expected) on the early Limburg velar realisations of West Germanic / $\hat{a}$ / and lengthened / $\check{a}$ /.

Coastal velarisation of secondary short *ǣ* in *bracht*, *zacht* is considered old (premediaeval) and has strongly regressed, whereas Brabantine velarisation of long/lengthened Wgm. *â*, *ǣ* (*staan*, *maken*) is thought to be more recent - at least post-mediaeval - and has a remarkable record of distribution into the surrounding dialect areas. I will show how the distribution pattern of the vowel type in *bracht* in the 14<sup>th</sup> is remarkably different than the one in the 20<sup>th</sup> century and that both of the abovementioned types of velarisation have to be taken into account to explain the difference.

## 5.2. EXISTING DATA FOR *BRACHT*

### 5.2.1. Middle Dutch

The fourteenth Century situation is visualised in map 2.<sup>37</sup> In the main map, the darker patches show the occurrences of <bracht>, whilst the brighter ones have concentrations of <brocht>. In the inlay, the darker patches show the areas with the most data, so we have to focus on those areas as the findings in the brighter areas are too sparse to be reliable. The colouring of these areas on the main map is based on an extrapolation of the surrounding data.

This mediaeval dialect data, as rendered through the 14th century charter corpus, reveals that <brocht>-orthographies occurred in the old western counties of Flanders and Holland/Utrecht. To the east, <bracht> - spelled with an <a> - prevails in Brabant, Limburg and in the North East. We think we can safely assume that <brocht> represents a more velar pronunciation than <bracht>.

The East of Flanders, and particularly the Dender-area, proves to be some sort of transition zone, where both variants occur. Brussels, however, is clearly situated in the <bracht>-zone and has, itself, five <bracht>-occurrences.

---

<sup>37</sup> A map with 13<sup>th</sup> Century data by Berteloot (1984: II, 12) shows a similar distribution, albeit with fewer data.

### 5.2.2. The current dialects

Phonological data from the 20<sup>th</sup> century shows a different distribution of open/medial and velar pronunciation. Most of the Dutch dialects now have some sort of short, velar vowel, ranging from [ɔ] to [o] and even to [u] in some places.<sup>38</sup> Map 3 is a rendering of modern phonological GTRP-data.<sup>39</sup>

The bright patches on the main map represent a vast area where the velar vowel occurs: the entire Brabantine area - including Brussels - has this vowel type. Open/medial vowels occur mostly in the Eastern dialects, and, to a lesser extent, in part of the Holland/Utrecht area.<sup>40</sup>

### 5.3. WHAT POTENTIALLY HAPPENED

The observer of these two maps might be fooled in thinking that ingvaeonic velar realisation of /ɑ/ (i.e. /ɔ/) in *bracht* has expanded dramatically towards the east since the 14<sup>th</sup> Century. This would, however, conflict with the socio-cultural status of the areas involved. Brabant, an economic powerhouse in the 15<sup>th</sup> and 16<sup>th</sup> Centuries, is generally considered to be the epicentre of linguistic renewal of that time, rather than a peripheral area at the receiving end. A hypothesis that takes this into account, has already been provided in FAND (I: 16): Brabantine *brocht* is endogenic, and it finds its origins in a more recently long or lengthened main vowel, that would then be receptive to Brabantine long /a:/ velarisation. This observation has already been made by Van Loon (1986: 42-44). The combined process would have generated the following (consecutive) output: br[ɑ]cht > br[a:]cht > br[ɔ:]cht > br[ɔ]cht. As no sign of this latter process can be found in the 14<sup>th</sup> Century data for Brabant, we can only assume that in the following couple of centuries, Brabantine occurrences of velar vocalism in *bracht* can be found. This is where the Brussels diary of Jan De Pottre comes in.

---

<sup>38</sup> Some places to the east and south of the city of Antwerp.

<sup>39</sup> A more detailed phonological map of the vowel in *bracht* can be found in FAND (I: 9).

<sup>40</sup> Introduction of (exogenic) bracht in Holland might be due to the standardisation process of the Dutch language.

#### 5.4. BRUSSELS DATA FROM THE 16TH CENTURY: JAN DE POTTRE (°1525-<sup>+</sup>1601)

I only excerpted the part of the text that accounts for De Pottre's every day life (rather than the first part that seems to contain second-hand material on previous years).<sup>41</sup>

I found a total of 33 occurrences of the lexemes *bracht*, *brachten* and *gebracht* combined:

gebr<a>cht	4	gebr<o>cht	11
br<a>chten	0	br<o>chten	11
br<a>cht(e)	1	br<o>cht(e)	6
<i>total</i> <a>	5	<i>total</i> <o>	28

To me, it seems very unlikely that De Pottre would use the orthography <o> in *bracht* if it did not represent a somewhat velar vowel. Especially the fact that both orthographies (<bracht>/<brocht>)<sup>42</sup> occur, reveals that this was an issue because of which the author, maybe subconsciously, was struggling with the impact of his own speech on his writing. The <brocht> occurrences outnumber the <bracht> ones by far. This might mean that the velar vowel quality in *bracht* occurred in a Brabantine dialect, and in a pretty important one at that. This implies that the hypothetical trajectory (br[ɑ]cht > br[a:]cht > br[ɔ:]cht > br[ɔ]cht) was operational before the time that De Pottre put his thoughts to paper, which was in the second half of the 16<sup>th</sup> century.

---

<sup>41</sup> I used the 1861 edition by B. de St. Genois, as rendered on the DBNL website ([http://www.dbnl.nl/tekst/pott013dagb01\\_01/pott013dagb01\\_01\\_0003.htm](http://www.dbnl.nl/tekst/pott013dagb01_01/pott013dagb01_01_0003.htm)). A collated version of the original manuscript - which I intend to provide during the ongoing project - might yield slightly different result.

<sup>42</sup> In their representation, I strip the different attestations from all variety except the main vowel, so *bracht*, *brocht* here also represent *gebracht/gebrocht*, *brachte/brochte*, etc.

## 6. CONCLUSION

The velar realisations of the vowel in Dutch *bracht*, *gebracht*, etc, as they occur in many Dutch dialects now, have possibly been caused by (at least) two different phenomena:

1) Old ingvaeonic velarisation that occurred during a general Germanic process of lengthening and nasalisation caused by the old post-vowel consonant cluster nasal + fricative + stop: brǣŋχt > brē̃χt > brōχt > bröχt;

2) A younger, temporary lengthening of the vowel in *bracht*, co-occurring with post-mediaeval Brabantine velarisation of long /a:/. This is represented in the following trajectory: br[a]cht > br[a:]cht > br[ɔ:]cht > br[ɔ]cht.

The currently available dialect databases and charter corpora reveal that phenomenon 1 was most definitely completed by the 14<sup>th</sup> Century, when 2 had not started yet; whereas 2 is very prominent in the current state of affairs.

A quick browse through the 16<sup>th</sup> century Brussels diary of Jan De Pottre has narrowed down the timeframe for 2 to some time within the 15<sup>th</sup> and 16<sup>th</sup> Centuries. I am hopeful that a broad corpus of similar texts, compliant with certain criteria (original, autographical, etc.) will help us to pinpoint many more phenomena in this transitional era in the history of Dutch.

## References:

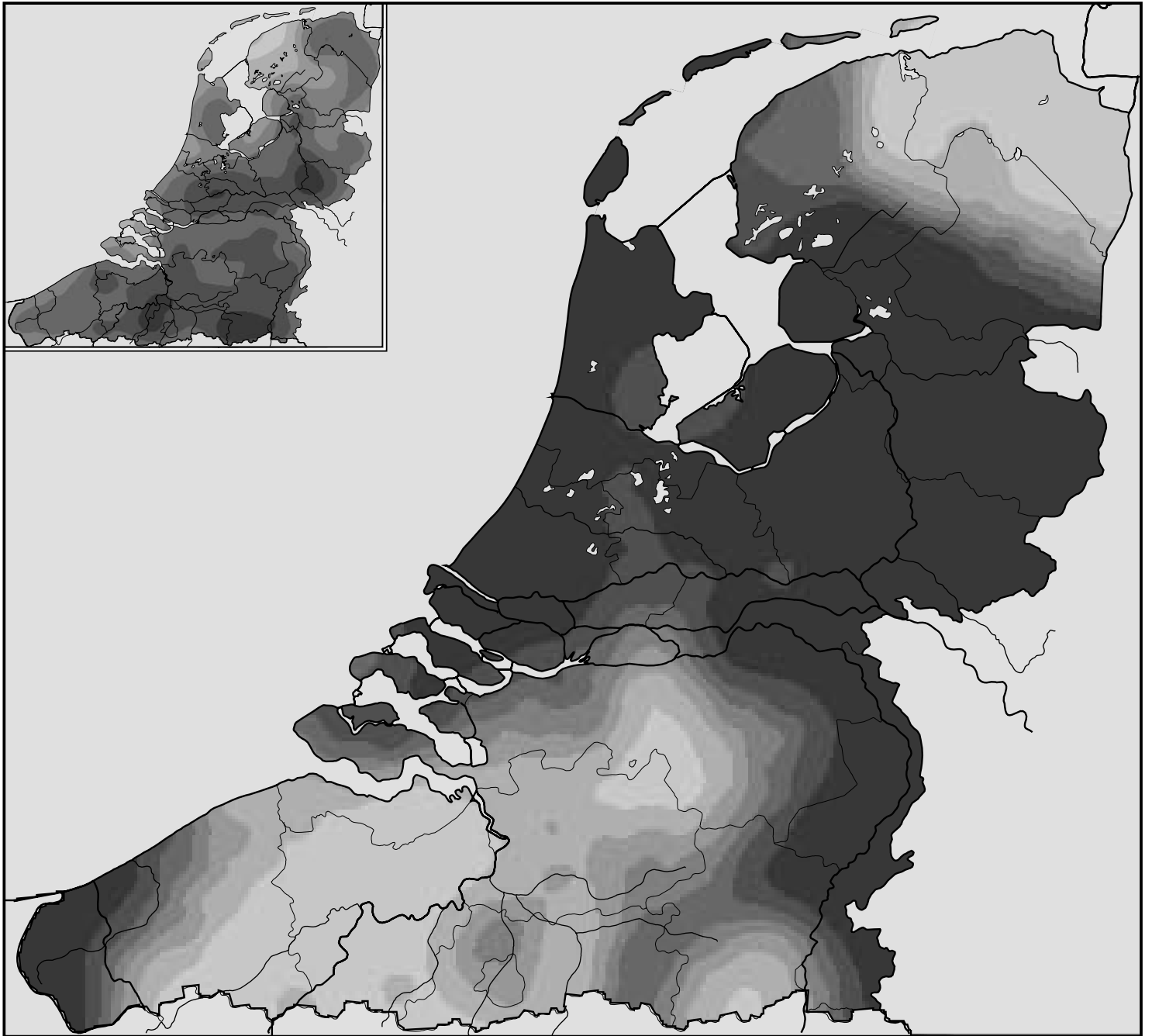
- Beaurepaire, P.-Y., D. Taurisson (éd.), 2003, *Actes du colloque international de Montpellier 'Les ego-documents à l'heure de l'électronique. Nouvelles approches des espaces et des réseaux relationnelles*. Montpellier.
- Berteloot, A., 1984, *Bijdrage tot een klankatlas van het dertiende-eeuwse Middelnederlands*. Gent.
- De Wulf, C., 2009 (to be expected), *Klankatlas van het veertiende-eeuwse Middelnederlands*. PhD University of Gent.
- FAND = Goossens, J., J. Taeldeman, G. Verleyen, 1998, *Fonologische Atlas van de Nederlandse Dialecten I*. Gent.
- Gysseling Corpus = Gysseling, M., 1977, *Corpus van Middelnederlandse teksten (tot en met het jaar 1300) uitgegeven door M. Gysseling. M.m.v. en van woordindices voorzien door W. Pijnenburg. Reeks I: Ambtelijke bescheiden*. 's-Gravenhage.
- Mooijaart, M.A., 1992, *Atlas van Vroegmiddelnederlandse Taalvarianten*. Utrecht.
- Rem, M., 2003, *De taal van de klerken uit de Hollandse grafelijke kanselarij (1300-1340). Naar een lokaliseringprocedure voor het veertiende-eeuwse Middelnederlands*. Amsterdam.
- Schönfeld 1970 = Schönfeld, M. & A. Van Loey, 1970<sup>8</sup>, *Schönfelds Historische grammatica van het Nederlands. Klankleer, vormleer, woordvorming*. Zutphen.
- Vanacker, F.V., 1963, *Syntaxis van gesproken taal te Aalst en in het Land van Aalst in de XVde, de XVIde en de XVIIde eeuw*. Brussel.
- Van Loon, J., 1986, *Historische fonologie van het Nederlands*. Leuven.
- Van Loon, J., 2001, Op zoek naar de ideale tekst. In: *Verslagen en mededelingen van de Koninklijke Academie voor Nederlandse Taal- en Letterkunde*, 111 (2001), p. 413-430.
- Van Reenen, P., 2002, Het 14de-eeuwse Middelnederlandse oorkondencorpus als dynamisch-systematisch referentiekader voor taalkundig onderzoek. In: Vanhoutte, E. (red.), *Talig Erfgoed. De zuidelijke Nederlanden in de 14de eeuw*. Gent, p. 59-72.
- Van Reenen, P. & M. Mulder, 1993, Een gegevensbank van 14de-eeuwse Middelnederlandse dialecten op computer. In: *Lexikos 3 (1993)*, p. 259-279.
- Van Rossem, C. & H. Van der Voort, 1996, *Die Creol taal. 250 years of Negerhollands texts*. Amsterdam.
- Wattel, E. & P. Van Reenen, 1995, Visualisation of extrapolated social-geographical data, Structures and contingencies in computerized historical Research. In: O. Boonstra et al, *Cahier VGI 9, structure sand contingencies in computerized historical research. Proceedings of the IX International Conference of the Association for History & Computing, Nijmegen 1994*. Hilversum, p, 253-262.

## Summary

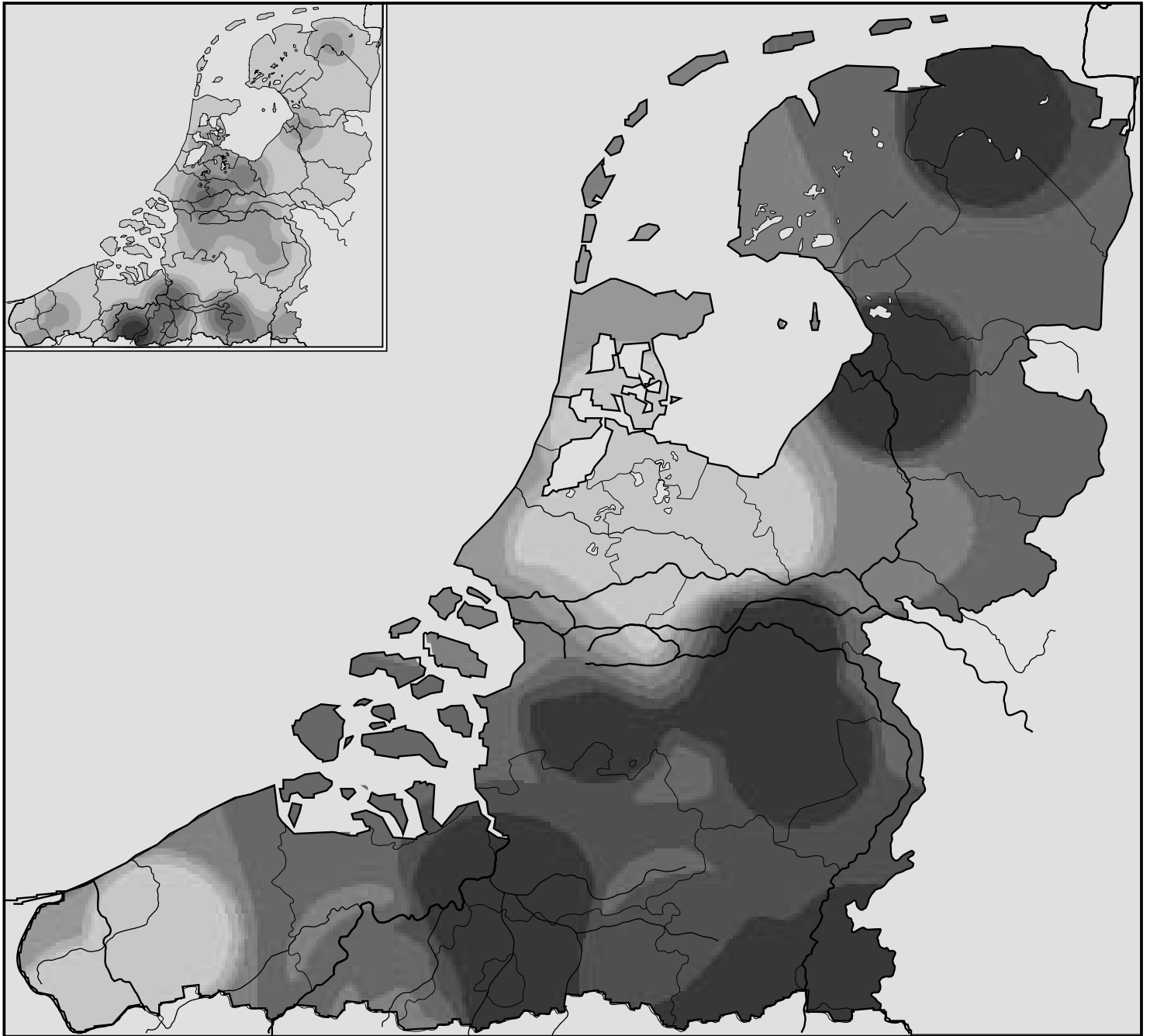
In 2008 the Royal Academy for Dutch Language and Literature (KANTL) has commissioned its research cell, the Centre for Scholarly Editing and Document Studies (CTB), to build a corpus containing 15<sup>th</sup> and 16<sup>th</sup> century ego documents, hence aiming to extend existing mediaeval text corpora in time. In this paper I will discuss the existing 13<sup>th</sup> and 14<sup>th</sup> century corpora (1) and the need to extend them in time (2). Then I will give an overview of various similar initiatives for the subsequent periods (3). Finally, I will reveal how the CTB/KANTL plans to add to these during the current project (4). This means that I will transfer selection criteria of other types of non-literary texts (charters) to autobiographical text (so-called *ego documents*). I will also discuss some examples of manuscripts that are eligible for this kind of research according to these selection criteria. A modest case study involving one of these ego documents will conclude this paper (5).



Map 1: 20th century m[a.]ken/m[o.]ken



Map 2: 14th century <br><br>



Map 3: 20th century br[a]cht/br[o]cht

